

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/122466>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

Categorized Question Template Generation for Ontology-based Assessment Questions

Noor H. Ibrahim Teo, and Mike S. Joy

Abstract—This paper discusses how to generate question templates by interpreting existing assessment questions from textbooks and validating concepts from existing ontologies. Previous work shows that most of the concepts that exist in the ontologies can be used as keywords to generate useful assessment questions. In this paper, we consider questions generated from a variety of ontologies, and question taxonomies, which include definitions, concept completions, and comparisons. Evaluation of novel techniques for question generation provides insights on how to extract a pattern from assessment questions in a textbook in order to create question templates that can be used with concepts presented in ontologies to generate useful questions. This paper discusses the method used and the experimental results.

Index Terms— Question Template, question generation, domain ontology.

I. INTRODUCTION

Question categorization is a vital process in question answering systems. The role is to indicate the type of answer for each question. As knowledge is growing, the number of question created in a system becomes bigger. Some of the questions are repetitive in their patterns. For example, the system might have a significant number of definition questions that have the same question stem but vary in terms of the keyword. This type of question could be formalized to allow instantiation of other concepts. Question templates are used in order to enable the instantiation and reusability of question stems.

Question templates may have placeholders for the key terms as well as verb lemmas (*do*, *be*) [1]. Question generation has been used for education purposes in different ways such as developing knowledge/skills, knowledge assessment and Socratic dialogues [2]. Selection of appropriate question templates is crucial in order to design meaningful questions. Inappropriate selections of question words (such as *what*, *define*, *explain*) will also affect the meanings of the questions. For example in Table I, if we instantiate the question word (QW) in question 1, which is ‘*What*’ with another keyword like ‘*Differentiate*,’ it would yield an incorrect meaning for the question. The solution we

adopt is to classify the question according to the appropriate category of Graesser’s taxonomy that we believe would help in generating a correct instantiation of the question from the question template. Therefore, there is a need to design categorized question templates rather than using a general question template to ensure valid questions are generated by the question template instantiation process.

TABLE I: GENERAL QUESTION TEMPLATE

Num	Questions	Template
1	What is an Operating System?	[QW] is [C]
2	What is a Memory Management?	[QW] is an [C]
3	Define communication media.	[QW] [C]
4	Differentiate between segmentation and paging in Operating System.	[QW] between [C1] and [C2] in [C3]
5	Explain the circular wait in deadlock.	[QW] the [C1] in [C2]

[QW]: Question word, [C]: Concept

This section first reviews the literature on question categories in education. Then, it presents current work on Automatic Question Generation (AQG) using question templates and finally the question templates approach.

II. QUESTION CATEGORY

Questions play significant roles in learning such as to assessing knowledge, triggering more ideas on knowledge being learned, and keeping learner focus during the learning process. Due to various roles for the questions, different ways of asking questions are required. Questions used in tutoring have been widely studied by the researchers. Tutorial questions have been extensively explored [3], and their question classification, consisting of 22 categories of questions to accommodate difference level of knowledge, has become a point of reference for subsequent research. Further analysis has been done, and the 22 categories of question have been further classified into four common mechanisms which are a correction of knowledge deficit, monitoring common ground, social coordination of action and control of conversation and attention [1].

Following the question categories presented in [3], another question category known as identity has been suggested in [4]. This question type has been proposed to investigate how its question type relates to the process of clarifying the question asked during a tutoring session.

The question taxonomy proposed in [3] has been categorized further into five categories, which are descriptive, method, explanation, comparison and preference, and this is referred to as a *primary taxonomy* [5], and is claimed might be useful for generating tutoring question.

Manuscript received May 24, 2018. This work was supported in part by the University of Warwick and Universiti Teknologi MARA.

N. H. Ibrahim Teo is with the Department of Computer Science at the University of Warwick, Coventry, CV47AL, UK, on study leave from University Teknologi MARA, Jasin Melaka, Malaysia. (e-mail:shimateo@uitm.edu.my / N.H.Ibrahim-teo@warwick.ac.uk).

M. S. Joy is with the Department of Computer Science at the University of Warwick, Coventry, CV4 7AL, UK. (e-mail:m.s.joy@warwick.ac.uk)

III. PRIOR WORK ON QUESTION GENERATION USING QUESTION TEMPLATE

The Programming Language Online Exam Platform (PLOEP) was suggested in [6] to resolve plagiarism issues amongst students when writing source code. The authors claim that fixed exercises increase the chances of plagiarism, which then reduce the learning effect. The generated ontology model consists of a concept part that uses basic set theory and a C++ implementation part that has been tested and has been reported as to be able to produce 3,624 open questions from four question templates. The significant number of questions produced makes it possible to generate a dynamic question for each student and hence reduce the chances of plagiarism. However, this approach is expensive to be implemented in other domains, since it requires knowledge for connecting the C++ implementation with other domains. It also requires an expert to design appropriate question templates that can be manipulated and reused using knowledge presented in an ontology model.

Semantic annotations used for marking student answers for predefined test questions have been discussed in [7]. Annotations for a question are provided by the teacher together with generated questions. An answer annotation is a semi-automatically generated ontological element which is detected by Natural Language Processing (NLP) using an incremental knowledge acquisition algorithm. There are three steps involved in this algorithm for the process to map a linguistic expression to the knowledge entity, namely (i) preparing the sentence until it gets tags using POS Tagging, (ii) collecting search phrases that identify linguistic expressions representing an ontological element, and (iii) using grammatical patterns to put the knowledge expression into context.

Examples of annotation-based stems and strategies for generating MCQ distractors were introduced in [8], based on the assumption that the questions will become more difficult if the choices of distractor are similar. The system (SeMCQ) was developed as a plug-in for the Protégé ontology editor [9] and uses a domain ontology to generate MCQ items. The semantic similarity used the combination of a text similarity measure with an ontology element similarity measure. The authors use a ‘question template’ to mean the semantic interpretation that applies during the mapping process between the domain ontology and the target MCQ ontology. The predefined ‘question template’ was designed to follow the level of difficulty in Bloom’s Taxonomy. The ‘question template’ was populated with various concepts and annotations from the domain ontology. Predicate logic was used to represent the relationship of class and annotation to be used in the ‘question template’. However the relationship used in the domain ontology was restricted to inheritance or an ‘is-a’ relationship, and might not be able to represent specific course domains in detail. In addition, the ‘question template’ was quite restricted and might need to be enriched to produce more choices for questions.

IV. HOW TO FORM QUESTION TEMPLATES

This section discusses how to design a question template pattern by using an ontology as a source of knowledge. Four

main processes will be discussed in the following subsections A to D, in order to obtain question templates. Modified questions obtained from an Operating Systems textbook [10] were used. We apply a simple formalization of a sentence using instantiation by extracting nouns and replacing them with placeholders in the sentence, to obtain question templates.

A. Question Collection and Selection

Assessment questions for the subject to be evaluated can be collected from different sources such as textbooks, online material, quizzes, and tests. The selection of questions depends on the purpose of the questions to be generated, that is, for learning purposes or for tests.

B. Question Classification

Selected questions were classified into appropriate question categories using the coding scheme in [5]. Table II shows the examples of categorized questions.

TABLE II: SAMPLE OF CATEGORIZED QUESTION

Question	Category
What is an interrupt?	Definition
What are the four components of a computer system?	Feature specification
What is another term for multiprocessor system?	Concept completion
How does a clustered system differ from a multicore system?	Comparison
Differentiate between mechanism and policy in CPU scheduling.	Comparison
Explain the convoy effect in process scheduling.	Concept completion
Explain the monitor in concurrency control.	Concept completion
Discuss the following term in Deadlock: i. Detection ii. Recovery	Concept completion
What is the purpose of system program?	Concept completion
Explain the advantage of Shortest Job First algorithm	Concept completion
List five services provided by an operating system.	Feature specification
A spinlock is a type of mutex lock. True or False.	Verification
List three disk scheduling algorithm.	Feature specification
Name one type of file system.	Feature specification
What does the acronym VFS refer to?	Concept completion
Explain the term bandwidth.	Concept completion
Define context switching.	Definition

C. Removing Question with the Same Pattern

Questions with the same pattern are removed to allow unique question templates to be stored in the question template repository. ‘Same pattern’ in this context refers to a set of question such that all tokens in the sentences are matched except for the keywords. The following examples illustrate the approach.

Example 1:

What is Operating System?	Selected
What is Security?	Removed
What is Scheduling?	Removed

Example 2:

What is the difference between non pre-emptive and pre-emptive scheduling	Selected
---	----------

algorithm?

What is the difference between protection and security? Removed

What is the difference between unsafe and safe state? Removed

In Example 1, the three definition questions have the same pattern which is ‘What is’ and differ in their keywords which are *Operating System*; *Security*; *Scheduling*. While in Example 2, the comparison questions have a pattern ‘What is the different between ... and ...?’ but different keywords are substituted. Therefore, for each of the examples, one question will be selected, and the other two will be removed.

D. Removing Question Keyword and Question Word

Question keywords are replaced with a variable **X** to form question templates. For example:

Question: What is **Operating System**?

Question Template: What is **X**?

Question: What is the difference between **non pre-emptive** and **pre-emptive**.

Question template: What is the difference between **X** and **Y**?

The next step is to remove each question word, in order to create formalized question templates. Examples 1 and 2 are a sample question template and its formalized question template respectively. The acronym QWD will indicate Question Word for a Definition, and QWC indicates Question Word for Comparison. Examples are as follows:

Example 1:

Question Template:

What is X?

Define X?

What is meant by a/an X?

Formalized Template: **QWD X?**

Example 2:

Question template:

What is the difference between X and Y?

Describe the differences between X and Y?

Compare between X and Y?

Formalized Template: **QWC X and Y?**

V. EXPERIMENTAL SETUP

Operating System (OS) reviewed questions were collected from a textbook [10] consisting of 263 questions from various question types which are different in the depth of knowledge. In this work, the interest is to investigate descriptive and comparison types of questions thus only those types of questions were selected. For the purpose of this work, simple questions that contain no more than 3 keywords were selected.

VI. RESULT AND DISCUSSION

A. Result for Categorizing OS Reviewed Questions

TABLE III shows the results obtained after classifying 263

questions into question categories proposed by [3].

TABLE III: NUMBER OF QUESTIONS FOR EACH QUESTION CATEGORIES

QUESTION TAXONOMY	NUM. OF QUESTION
Definition	7
Comparison	4
Concept Completion	110
Verification	61
Feature Specification	40
Others	41

Most of the questions analyzed fall under the descriptive and comparison types of question. The descriptive questions include concept completion, feature specification, and definition.

The following form a sample of questions for each of the category.

Definition:

Define **page fault**.

What is meant by a **kernel thread**?

Comparison:

What is the difference between **volatile** and **non-volatile** storage?

How does a **clustered system** differ from a **multicore system**?

Concept Completion:

Explain what a **Page fault** is.

Explain the term “**Collision Detection**.”

Verification

True or False? A **physical address space** is at least as large as a **virtual address space**.

Features Specification

List the three general **disk scheduling algorithms**.

Provide two advantages of **multiprocessor systems**.

B. Questions That Can Be Generated From an Ontology

In this case, not all questions can be generated from ontologies because of the complexity of the questions. For example, the question “What part of the operating system made decisions with regards to which job will run?” could not be generated using an ontology. This type of question would need more complex question representation to represent the information in the question.

However, the simple question sentence could be generated using template-based question generation from the ontology. Table IV shows sample questions that can be generated using ontology.

TABLE IV: SAMPLE OF QUESTIONS THAT CAN BE GENERATED

Questions	Question templates
What is an interrupt?	What is an/a [C1]?
Explain the convoy effect in process scheduling.	Explain the [C1] in [C2].
Differentiate between mechanism and policy in CPU scheduling.	Differentiate between [C1] and [C2] in [C3].
Explain the term bandwidth.	Explain the term [C1].

C. Result for Question Template Formalization

The selection of the sample questions is discussed in [11]

and is processed further to form question templates by applying the steps discussed in section IV. Table V shows the question templates obtained.

TABLE V: QUESTION TEMPLATES

Question template	Formalize template
Definition	
Define C	[QWD] [C]
What are C1 and C2	[QWD] [C1] and [C2]
What is C1, C2, and C3	[QWD] [C1], [C2], and [C3]
Define the following terminology:	[QWD][C]
a. C1	a. [C1]
b. C2	b. [C2]
c. C3	c. [C3]
Concept Completion	
What is a C1 in C?	[QWCC] [C1] [C]
In the context of C, what is C1?	[QWCC][C] [C1]
Comparison	
What is the difference between C1 and C2	[QWC] [C1] and [C2]
With regard to C, explain the difference between C1 and C2	[QWC] [C],[C1], and [C2]
Explain the difference between the C1, C2, and C3 for C.	[QWC] [C1],[C2],[C3] for [C]

VII. CONCLUSION

This kind of research aims to benefit instructors by providing support and intelligent assistance for the automatic generation of questions. This paper has discussed how to generate question templates by interpreting existing assessment questions from a textbook and validating concepts from existing ontologies. The evaluation has led to a further exciting analysis, which is how to extract a pattern from assessment questions in a the textbook in order to create question templates that can be used with ontology concepts to generate useful questions. This paper also discusses how to formalize the questions to obtain the patterns for categorizing the question templates. An experiment has been conducted to analyze real-world assessment questions and transform them into question templates. The question templates obtained from this work will be useful for generating question using ontologies. In future work, the mapping strategies between question templates and concepts in ontologies will be explored.

REFERENCES

[1] A.M. Olney, A.C. Graesser, and N.K. Person (2012). Question Generation from Concept Maps. *Dialogue & Discourse* 17, 3(2). pp. 75-99.

[2] N.T. Le, N.P. Nguyen, K. Seta, and N. Pinkwart. (2014). Automatic Question Generation for Supporting Argumentation. *Vietnam Journal of Computer Science*, 1 pp. 117-127.

[3] A.C. Graesser, and N.K. Person. (1994). Question Asking During Tutoring. *American Educational Research Journal*, 31 pp. 104-137.

[4] D. Bodoff, and D. Raban. (2015). Question Types and Intermediary Elicitations. *Journal of the Association for Information Science and Technology*. 67(2). pp. 289-304.

[5] R.D. Nielsen, J. Buckingham, G. Knoll, B. Marsh, and L. Palen (2008). A taxonomy of questions for question generation. *Proceedings of the Workshop on the Question Generation Shared Task and Evaluation Challenge*.

[6] C.Y. Lin, and C.C.Chou (2011). Effectiveness of Ontology-Based Online Exam Platform for Programming Language Education. *Computing in Civil Engineering* pp. 907-914.

[7] D. Castellanos-Nieves, J.T. Fernández-Breis, R. Valencia-García, R. Martínez-Béjar, and M. Iniesta-Moreno. (2011). Semantic web technologies for supporting learning assessment. *Information Sciences* 181(9) pp. 1517-37.

[8] M. Cubric and M. Tosic (2011). Towards automatic generation of e-assessment using semantic web technologies. *International Journal of e-Assessment* 1(1).

[9] H. Knublauch (2014). "Ontology-driven software development in the context of the semantic web: An example scenario with Protege/OWL 1st International Workshop on the Model-Driven Semantic Web (MDSW2004) pp. 381-401.

[10] A. Silberschatz, P.B. Galvin, and G. Gagne (2012) Operating System Concepts. New Jersey: John Wiley & Sons.

[11] N.H. Ibrahim Teo and M.S. Joy (2016) Validation of Course Ontology Elements for Automatic Question Generation. In: *Vincenti G., Bucciero A., Helfert M., Glowatz M. (eds) E-Learning, E-Education, and Online Training. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 180.



Noor H. Ibrahim Teo obtained her BA in information technology and MA in computer science at Universiti Teknologi MARA, Malaysia, and works as lecturer at the Department of Computer Science and Mathematics, Universiti Teknologi MARA, Malaysia.

She is currently in her final year of her PhD in computer science studies at University of Warwick, United Kingdom. Her research interests center around computer science and education technology. Her current research is ontologies for automatic question generation.



Mike Joy received the MA degree in Mathematics from Cambridge University, the MA degree in postcompulsory education from the University of Warwick, and the PhD degree in computer science from the University of East Anglia. He is currently Professor at the University of Warwick. His research interests focus on education technology and computer science education.